Iranian Food Science and Technology Research Journal
Vol. 16, No. 4, Oct- Nov 2020, p. 479-491

نشریه پژوهشهای علوم و صنایع غذایی ایران
جلد 16، شماره 4، مهر - آبان 1399، ص. 491-479

مقاله علمی-پژوهشی

پیاده‌سازی یک سامانه ماشین بویایی برای طبقه‌بندی انواع مختلف فلفل سیاه براساس منشاء جغرافیایی و تشخیص تقلب در فلفل سیاه هندی

فائزه جمالیزاده[1]- مهدی قاسمی ورنامخواستی[2*]- مهدی قاسمی نافچی[3]- مجتبی توحیدی[4]- مجید دولتی[5]

تاریخ دریافت: 1397/08/15
تاریخ پذیرش: 1398/06/17

چکیده

ادویه‌جات از با ارزش‌ترین گیاهان دارویی مورداستفاده در صنایع غذایی و علم پزشکی هستند و با توجه به تفاوت کیفیت و قیمت بین گونه‌های مختلف، تشخیص، طبقه‌بندی و جداسازی آن‌ها براساس خلوص و درجه کیفیت از اهمیت بالایی برخوردار است. ادویه‌ها درکشورهای مختلفی از جمله هندوستان، پاکستان، چین و کشورهای آسیای شرقی و جنوبی تولید می‌شوند. در این پژوهش، یک سامانه ماشین بویایی بر پایه هشت حسگر نیمه هادی اکسید فلزی در ترکیب با روش‌های تشخیص الگو به‌منظور طبقه‌بندی و جداسازی ادویه فلفل سیاه براساس منشاء جغرافیایی و تشخیص تقلب‌های کبابه چینی و پودر هسته خرما به‌کارگرفته شد. به‌منظور تحلیل داده‌های استخراج شده از سیگنال پاسخ حسگرها از روش تحلیل مولفه‌های اصلی (PCA) استفاده شد. براساس نتایج حاصل، آنالیز مولفه‌های اصلی با مجموع دو مولفه اصلی اول 96% برای نمونه‌های فلفل سیاه براساس منشاء جغرافیایی و 95% برای تقلب‌های کبابه چینی و هسته خرما از واریانس داده‌ها قابل توصیف است. همچنین از سه روش تحلیل تفکیک خطی (LDA)، ماشین بردار پشتیبان (SVM) و درخت تصمیم‌گیری (DT) برای طبقه‌بندی نمونه‌ها استفاده شد. استفاده از روش LDA، برای نمونه‌های فلفل سیاه دقت طبقه‌بندی 100% و برای تقلب‌ها دقت 97/14% را نشان داد. نتایج نشان داد که SVM با تابع گاوسی بالاترین دقت را در طبقه‌بندی نمونه‌های فلفل سیاه، تقلب کبابه چینی و تقلب هسته خرما را دارد. همچنین میزان موفقیت روش DT در تفکیک و طبقه‌بندی نمونه‌های فلفل سیاه، تقلب کبابه چینی و تقلب هسته خرما به‌ترتیب 96/66% و 88/5% برآورد شد.

واژه‌های کلیدی: ماشین بویایی، فلفل سیاه، کبابه چینی، هسته خرما.

مقدمه

بخشی از گیاه معطر خشک‌شده که به‌صورت کامل یا پودرشده در آشپزی به‌عنوان طعم‌دهنده و رنگ‌دهنده استفاده می‌شود، ادویه نام دارد (Peter, 2012). ادویه‌ای مانند فلفل سیاه با نام علمی *Piper nigrum* در ردیف مهم‌ترین کالاهایی می‌باشد که بشر از دیرباز به تجارت آن اشتغال داشته است و کالایی گران‌قیمت در بازرگانی قدیم محسوب می‌شود.

کنترل کیفیت، مجموعه‌ای از فعالیت‌ها و تکنیک‌های اجرایی است که برای دستیابی به مشخصات کیفیت تعیین شده مورد استفاده قرار گرفته‌است و برای رفع علل عملکرد غیررضایت بخش در تمامی مراحل چرخه کیفیت به‌منظور نیل به‌کارآرایی اقتصادی انجام می‌گیرد (امیدبیگی، 1384). یکی از چالش‌های مطرح شده در مورد ارزیابی کیفی ادویه‌ها، تشخیص تقلب و ارزیابی میزان خلوص آن‌ها است. چرا که جعل و یا تقلب در موادغذایی، با انگیزه افزایش سود اقتصادی صورت می‌گیرد که تهدیدی بزرگ برای سلامت عمومی می‌باشد. تقلب هم می‌تواند به‌صورت اضافه کردن مواد ارزان از قبیل رنگ و حجم دهنده‌ها به ادویه‌ها صورت پذیرد و هم به‌صورت تقلب بر اساس منشاء جغرافیایی باشد. به‌طور مثال ادویه‌هایی از چین و پاکستان که از کیفیت پائین‌تری برخوردار هستند به جای ادویه‌های هندی به فروش می‌رسند (Banerjee *et al*., 2017; Gliszczyńska-Świgło *et al*., 2017).

علاوه بر تقلب‌هایی که براساس منشاء جغرافیایی صورت می‌گیرد، تقلب‌های مختلفی نیز به دلیل ارزش اقتصادی فلفل سیاه مرغوب (با منشاء جغرافیایی هندی) در این محصول انجام می‌شود. از متداول‌ترین تقلب‌های رایج در این ادویه با ارزش افزودن کبابه چینی (*Cubeb pepper*) و پودر هسته خرما می‌باشد که این عمل با عدم رضایت مصرف‌کنندگان همراه است. لذا تشخیص تقلب در فلفل سیاه

1، 2 و 4- به‌ترتیب دانشجوی کارشناسی ارشد، دانشیار و دکتری، گروه بیوسیستم، دانشکده کشاورزی دانشگاه شهرکرد.
3- استادیار، گروه مهندسی علوم باغبانی، دانشکده کشاورزی، دانشگاه شهرکرد.
5- استادیار، گروه بیوسیستم، دانشکده فنی و منابع طبیعی تویسرکان، دانشگاه بوعلی سینا همدان.
(* - نویسنده مسئول: Email: ghasemymahdi@ut.ac.ir)
DOI: 10.22067/ifstrj.v16i4.76455

به‌عنوان ادویه‌ای با ارزش در صنایع غذایی و علم پزشکی از اهمیت بالایی برخودار است.

تاکنون جهت بررسی کنترل کیفیت ادویه‌ها از روش‌های مختلفی مانند تست پنل[1](Li *et al*., 2013)، روش‌های آزمایشگاهی مانند کروماتوگرافی گازی (GC)[2] (Shafiqul Islam *et al*., 2006)، ماشین بینایی[3] (سیدی و عابدینی؛ 1390) و ماشین چشایی[4] (Baldwin *et al*., 2011) استفاده شده است. که بیشتر این روش‌ها پیچیده، وقت‌گیر و پرهزینه بوده و به افراد متخصص و ابزارهای دقیق و گران‌قیمت نیاز دارند. ماشین بویایی (بینی الکترونیکی)[5] ، یکی از ابزارهای نوین، مناسب، کارآمد و کم‌هزینه در سنجش کیفیت و طبقه‌بندی مواد غذایی است. این سیستم با استفاده از یک آرایه حسگری متشکل از حسگرهای گازی کار سیستم بویایی انسان را شبیه‌سازی کرده و به نوعی گازهای فرار متصاعد شده از مواد غذایی را بو می‌کند. حسگرهای گازی مورداستفاده کمیت‌های شیمیایی را به کمیت‌های الکتریکی تبدیل می‌کنند. سیگنال‌های به‌دست آمده از تغییرات بوی مواد به واحد پردازشگر ارسال شده و با توجه به تغییرات در روند سیگنال‌ها، صفت مورد نظر استخراج شده و در ادامه با استفاده از روش‌های تشخیص الگو تغییرات آن به‌دست می‌آید (قاسمی ورنامخواستی، 2011; Kiani *et al*., 2016b; Bhattacharyya *et al*., 2008). استفاده از ماشین بویایی در چند سال اخیر به‌منظور ارزیابی اصالت محصولات و تقلب صورت گرفته در آن‌ها افزایش چشمگیری داشته است. از جمله این پژوهش‌ها می‌توان به زعفران (Heidarbeigi *et al*., 2015)، ادویه مخلوط (Banach *et al*.,2012) ، فلفل سفید (Liu *et al*., 2013) اشاره کرد. بنابراین هدف پژوهش حاضر، بررسی توانایی سامانه ماشین بویایی در شناسایی و طبقه‌بندی ادویه فلفل سیاه براساس منشاء جغرافیایی و همچنین تشخیص تقلب در نمونه هندی این ادویه می‌باشد.

مواد و روش‌ها

ابتدا نمونه‌های فلفل سیاه براساس منشاء جغرافیایی (هندی، پاکستانی و چینی)، کبابه چینی و پودر هسته خرما از نمونه‌های موجود در بازار به‌صورت فله و پودر نشده (برای حصول اطمینان از کیفیت) تهیه شدند سپس نمونه‌ها آسیاب، و پودر شدند. مشخصات نمونه‌های آماده شده در جدول 1 آورده شده است. برای تهیه نمونه‌های فلفل سیاه تقلبی با درصدهای مختلف تقلب (وزنی)، یک ساعت پیش از انجام آزمایش‌ها با بینی الکترونیکی، کبابه چینی و پودر هسته خرما با نسبت‌های 10، 20 و 30% به فلفل سیاه هندی افزوده شدند. برای هر آزمایش 2 گرم از نمونه‌های آماده شده در داخل یک پتری دیش تمیز و فاقد رطوبت در محفظه مخصوص نمونه‌گیری ریخته شد. برای هر گروه نمونه‌ها، آزمایش‌ها ده بار تکرار گردید.

جدول 1- مشخصات نمونه‌های آماده شده

نمونه	نام انگلیسی	نام علمی	خانواده
فلفل سیاه	Black pepper	Piper nigrum	Piperaceae
کبابه چینی	Cubeb pepper	Piper cubeba L	Piperaceae
هسته خرما	Palm kernel	----	----

ماشین بویایی (بینی الکترونیک)

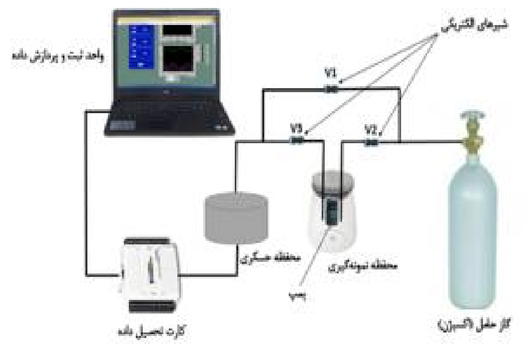
جهت تشخیص و طبقه‌بندی نمونه‌ها از سامانه ماشین بویایی دارای حسگرهای نیمه‌هادی اکسید فلزی (MOS) ساخته شده در دانشگاه شهرکرد استفاده شد (توحیدی و همکاران، 1395). این حسگرها از حساسیت و پایداری شیمیایی بالایی برخوردار بوده و در عین حال کم‌هزینه هستند و قادرند که یک کمیت شیمیایی را به یک سیگنال الکتریکی تبدیل کنند (کیانی و مینایی، 1394). سامانه ماشین بویایی از آرایه‌های حسگرها، محفظه حسگری و نمونه، سامانه تحصیل داده، منبع تغذیه، شیرهای برقی، پمپ و کپسول اکسیژن تشکیل شده است و سامانه ماشین بویایی مورد استفاده در این پژوهش در شکل 1 آورده شده است. مجموعه حسگرها شامل ترکیبی از 8 حسگر نیمه‌هادی اکسید فلزی با نام‌های تجاری MQ (Hanwai, China)، FIS (Osaka, Japan) و TGS (Glenview, USA) که هرکدام از آن‌ها به ترکیبات خاصی از مواد فرار نمونه‌ها واکنش نشان می‌دهند (Sanaeifar *et al*., 2015b). حسگرهای استفاده شده و مشخصات آن‌ها در جدول 2 نشان داده شده است.

به‌منظور رساندن دمای سطح حسگرها به مقدار مناسب (300- 500℃) ولتاژ 5 ولت که توسط کارخانه‌های سازنده توصیه شده است به گرم‌کن حسگرها داده شد. جهت اطمینان از پایداری دمایی حسگرها و بهترین عملکرد در محدوده دمایی مناسب، 30 دقیقه قبل از شروع آزمایش‌ها، مجموعه حسگرها توسط یک بورد واسط شروع به‌کار کرده تا به دمای مورد نظر برسد (توحیدی و همکاران، 1395). نمونه‌های آماده شده، 30 دقیقه قبل از شروع هر آزمایش‌ها و فرایند نمونه‌برداری به‌منظور تولید مواد فرار و پر شدن فضای بالای نمونه در محفظه نمونه‌گیری کاملا عایق قرار داده شدند. سامانه ماشین بویایی برای سه فاز تصحیح خط مبنا، تزریق بوی نمونه و اندازه‌گیری و پاکسازی حسگرها و محفظه نمونه برنامه‌ریزی شده است. در مرحله تصحیح خط مبنا به‌منظور پاکسازی

---

1 Panel Test
2 Gas Chromatography-Mass Spectrometry
3 Machine Vision
4 Electronic Tongue
5 Electronic Nose

محفظه حسگرها و رساندن پاسخ حسگرها به خط مبنا، گاز حامل (اکسیژن) به مدت 180 ثانیه از روی حسگرها عبور داده شد. در مرحله تزریق بو و اندازه‌گیری، فضای هد توسط پمپی که در محفظه نمونه تعبیه شده است، با دبی 1/3 لیتر بر دقیقه به محفظه حسگرها منتقل شد. زمان مورد نیاز برای رسیدن پاسخ حسگرها به مقدار ماکزیمم، 240 ثانیه در نظر گرفته شد. در این مرحله ولتاژ خروجی هر حسگر با توجه به نوع و میزان حساسیت آن، تغییر می‌کند. در مرحله پاکسازی حسگرها و محفظه، گاز اکسیژن به مدت 155 ثانیه از روی حسگرها عبور داده شد تا اثری از بو باقی نماند. زمان‌بندی سامانه برای هر نوع کاربردی منحصر به فرد است و با تغییر کاربری سامانه، هریک از مراحل فوق باید مجددا زمان‌بندی شوند (قاسمی ورنامخواستی، 2011). پاسخ حسگرها در مراحل ذکر شده به‌وسیله یک سامانه تحصیل داده برنامه‌ریزی شده با نرم‌افزار Lab View 2013 جمع‌آوری و ذخیره شد.



الف



ب

شکل 1- الف) طرح‌واره‌ای از سامانه ماشین بویایی مورداستفاده ب) سامانه ماشین بویایی مورداستفاده

جدول 2- نام، مشخصات و محدوده تشخیص آرایه حسگری بینی الکترونیک

نام حسگر	کاربردهای اصلی	محدوده تشخیص(ppm)
TGS813	متان، پروپان، بوتان	10000-500
MQ8	هیدروژن	10000-100
MQ4	متان، گاز طبیعی	10000-200
TGS822	بخار حلال‌های آلی	5000-50
TGS2620	الکل، بخار حلال‌های آلی	5000-50
MQ135	آمونیاک، الکل، دود، بنزن و دی‌اکسید کربن	300-10
MQ3	الکل	10-0/05
FIS	کنترل کیفیت هوا	10-1 (سولفید هیدروژن) 100-10 (اتانول، آمونیاک، هیدروژن)

استخراج ویژگی

ابتدا داده‌ها باید پیش‌پردازش شوند تا به شکل قابل استفاده در نرم‌افزارهای تحلیل تشخیص الگو درآیند. پیش‌پردازش داده‌ها وابسته به نوع حسگرها بوده و تاثیر بسیار زیادی روی عملکرد روش‌های تشخیص الگو دارد (Pearce *et al.*, 2003). پیش‌پردازش شامل سه مرحله تصحیح خط مبنا، فشرده‌سازی و نرمال‌سازی داده‌ها است. هدف از تصحیح پاسخ حسگرها با توجه به خط مبنا به‌منظور جبران انحراف و افزایش قدرت تشخیص حسگرها است (حاجی‌نژاد و همکاران، 1396). سه روش پیش پردازش اختلافی[1]، کسری[2]، نسبی[3]، بدین منظور وجود دارد (Kiani *et al.*, 2016a) که در این پژوهش روش کسری به‌کار رفته است. در این روش ((0) X) کمترین پاسخ حسگر پیش از مرحله اندازه‌گیری (خط مبنا) از پاسخ حسگری ((X (t)) در زمان t کم می‌شود و سپس نتیجه بر پاسخ حسگر تقسیم می‌شود، پاسخ پیش‌پردازش شده ((Y (t)) بی‌بعد و نرمال شده است و

1 Differential
2 Fractional
3 Relative

پاسخ حسگرهایی که سطح پاسخ بالا و یا پایینی دارند را جبران می‌کند (Arshak *et al*., 2004).

تجزیه و تحلیل آماری داده‌ها

تحلیل مولفه‌های اصلی (PCA[1])

تحلیل مولفه‌های اصلی روشی چندمتغیره و بدون ناظر است که در فشرده‌سازی داده‌های خطی و استخراج ویژگی‌ها استفاده می‌شود. PCA با نادیده گرفتن تفاوت بین کلاس‌ها و حفظ مولفه‌هایی از مجموعه داده‌ها که بیشترین تاثیر را در واریانس دارند برای کاهش ابعاد داده می‌تواند مورد استفاده قرار گیرد (Li *et al*., 2007). روش PCA با مشخص کردن الگوها و طبقه‌بندی داده‌ها به‌نحوی داده‌ها را بیان می‌کند که شباهت‌ها و تفاوت‌های آن‌ها مشخص‌تر شود (Mahmoudi, 2009) و در شرایطی که تعداد داده‌ها کم بوده و نمونه‌برداری از داده‌ها به‌صورت غیریکنواخت صورت گرفته باشد استفاده می‌شود. نتایج این روش توسط نمودارهای گرافیکی (اسکور[2] و لودینگ[3]) بیان می‌شوند.

LDA[4]

تحلیل تفکیک خطی یک روش طبقه‌بندی با نظارت می‌باشد که برای دسته‌بندی و طبقه‌بندی داده‌ها، کلاس‌ها را در نظر می‌گیرد و براساس تفاوت کلاس‌ها ویژگی‌های مهم را استخراج می‌کند. هدف این روش این است که داده‌های مربوط به یک کلاس تا حد امکان به هم نزدیک و داده‌های کلاس‌های مختلف از هم دور باشند. به عبارت دیگر روش LDA با به حداکثر رساندن پراکندگی بین کلاس داده‌ها و به حداقل رساندن پراکندگی درون کلاس‌ها به تفکیک بیشتر بین گروه‌های مختلف کمک می‌کند (Tudu *et al*., 2008). تحلیل تفکیک خطی نیز مانند تحلیل مولفه‌های اصلی در کاستن تعداد بعد داده‌ها و در شرایطی که تعداد داده‌ها زیاد باشد، استفاده می‌شود.

SVM[5]

ماشین‌بردار پشتیبان یکی از روش‌های یادگیری ماشینی و روشی با نظارت است. به‌طور کلی این روش برای طبقه‌بندی دو یا چند گروه و رگرسیون مورداستفاده قرار می‌گیرد. مبنای کاری دسته‌بندی روش SVM، دسته‌بندی خطی داده‌ها بوده و در تقسیم خطی داده‌ها سعی می‌شود خطی انتخاب شود که حاشیه اطمینان بیشتری داشته باشد. برای طبقه‌بندی داده‌ها از دو روش C- SVM و Nu- SVM استفاده می‌شود که مبتنی بر راه‌های متفاوت برای به حداقل رساندن تابع خطا می‌باشند. C در روش C-SVM براساس اطلاع از نویز موجود در داده‌ها انتخاب می‌شود و پارامتر Nu در روش Nu-SVM به‌عنوان مرز بالایی از خطاها و حد پایین برای بردار پشتیبان عمل می‌کند (Sanaeifar *et al*., 2014).

روش درخت تصمیم (DT[6])

درخت تصمیم‌گیری ابزاری قوی، ساده و قابل فهم برای پیش‌بینی و طبقه‌بندی داده‌ها است که با تفسیر آسان نتایج و رشد داده کاوی استفاده از آن در حال افزایش است (طلوعی اشلقی و همکاران، 1391). این روش برای انواع داده‌های عددی و طبقه‌ای قابل استفاده بوده و احتیاجی به محاسبات پیچیده برای دسته‌بندی داده‌ها ندارد. درخت تصمیم متشکل از تعدادی گره و برگ می‌باشد. برگ‌ها نشان دهنده کلاس‌ها هستند و در هر یک از گره‌ها با توجه به یک یا چند صفت تصمیم‌گیری صورت می‌گیرد. درخت ایجاد شده تا جایی که امکان جداسازی بهتر داده‌ها وجود نداشته باشد، بزرگ می‌شود. از پر کاربردترین ساختارهای تصمیم می‌توان از الگوریتم پر کاربرد C4.5 نام برد که در پژوهش‌های زیادی استفاده شده است. این الگوریتم در نرم‌افزار داده کاوی Weka، J48 نام دارد، این الگوریتم به‌صورت گرافیکی یک درخت تصمیم‌گیری ایجاد می‌کند که تفسیر آن از سایر روش‌های طبقه‌بندی ساده‌تر است.

پارامترهای زیر به‌منظور ارزیابی عملکرد درخت تصمیم استفاده شده است (Patil & Sherekar, 2013):

- TP Rate (True Positive): میزان دسته‌بندی درست داده‌ها را نشان می‌دهد.
- FP Rate (False Positive): میزان دسته‌بندی نادرست داده‌ها را نشان می‌دهد.
- Recall: نسبت میزان کل مشاهدات طبقه‌بندی شده و مرتبط در هر کلاس را به کل مشاهدات مرتبط بیان می‌کند.
- Precision: نسبت میزان کل مشاهدات طبقه‌بندی شده و مرتبط در هر کلاس را به کل مشاهدات بیان می‌کند.
- F-Measure : از رابطه ذیل به‌دست می‌آید:

$$0\mathrm{F-Measure}=\frac{2\times\mathrm{Recall}\times\mathrm{Precision}}{\mathrm{Recall}+\mathrm{Precision}} \qquad (1)$$

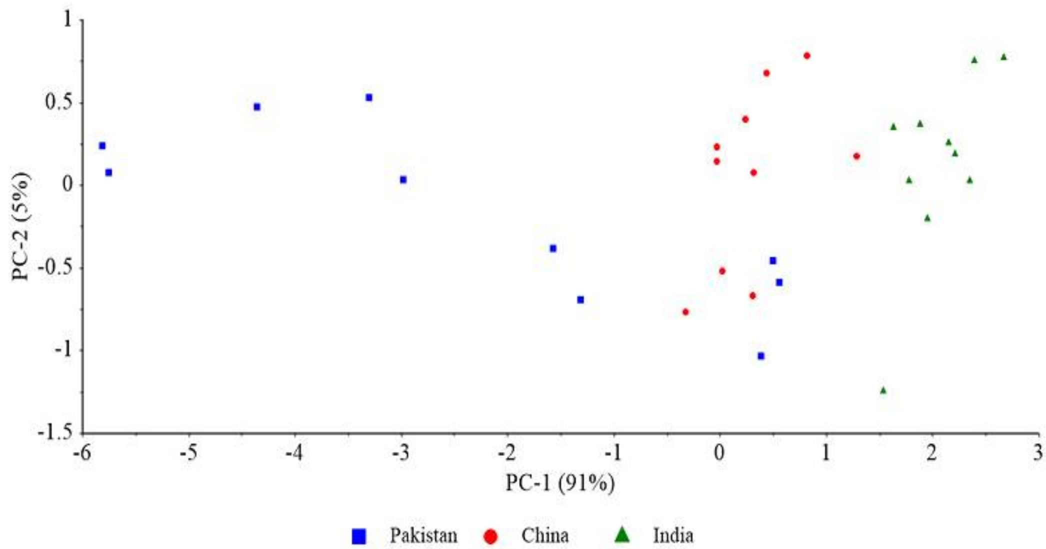
نتایج و بحث

نتایج آنالیز به روش PCA به کمک نمودارهای اسکور و لودینگ مورد بررسی قرار گرفتند. در شکل 2- الف نمودار اسکور دو مولفه اصلی برای نمونه‌های فلفل سیاه هندی، پاکستانی و چینی نشان داده

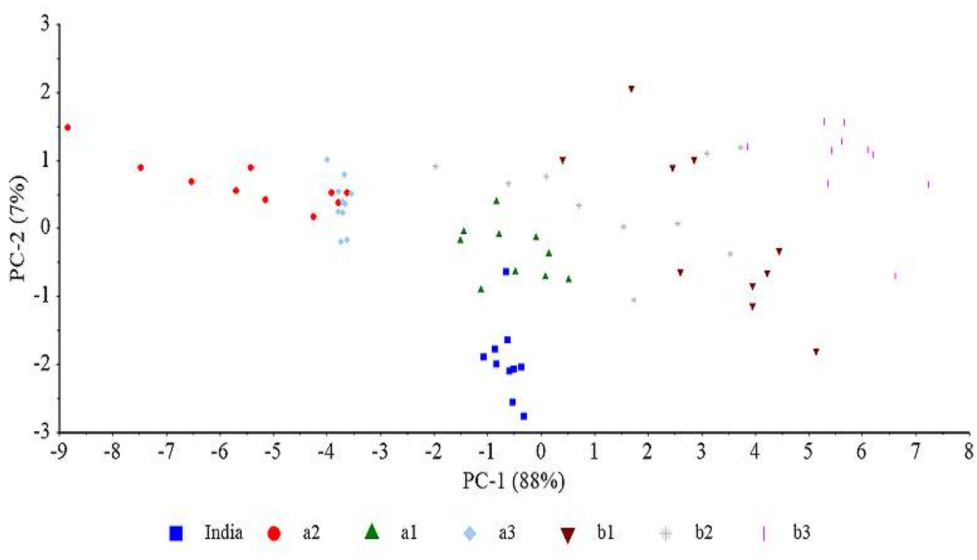
---

1 Principal Component Analysis
2 Score
3 Loading
4 Linear Discriminant Analysis
5 Support Vector Machine
6 Decision Tree

شده است. براساس نتایج، تمایز مناسبی بین نمونه‌های هندی با دو نمونه دیگر وجود دارد (Heidarbeigi *et al.*, 2015). دو مولفه اصلی PC1 و PC2 در این نمودارها بیانگر بیشترین مقدار واریانس داده‌ها می‌باشند که به‌ترتیب 91% و 5% واریانس بین نمونه‌ها و در مجموع 96% واریانس کل داده‌ها را توصیف کردند. نمونه‌های چینی به مقدار بسیار کمی با نمونه‌های پاکستانی همپوشانی دارند و بیانگر این مطلب است که پاسخ حسگرها به نمونه‌های چینی تا حدودی مشابه نمونه‌های پاکستانی است. همچنین شکل 2- ب نمودار اسکور مولفه‌های اصلی مربوط به سطوح مختلف تقلب‌های کبابه چینی و هسته خرما است. دو مولفه اصلی PC1 و PC2 به‌ترتیب 88% و 7% واریانس بین نمونه‌ها و در مجموع با 95% واریانس کل را توصیف کردند. با توجه به شکل نمونه هندی فلفل سیاه یا به عبارتی نمونه شاهد (بدون تقلب) تقریبا به‌صورت کامل از سطوح تقلب متمایز می‌باشد. اما سطح تقلب 20 و 30% در تقلب کبابه چینی با یکدیگر همپوشانی دارند. همچنین در تقلب هسته خرما بین سطوح تقلب 10 و 20% تمایز مناسبی وجود ندارد و قابل تفکیک از یکدیگر نیستند. اما سطح 30% به‌صورت کامل از نمونه شاهد و سایر سطوح تقلب مجزا است و با افزایش حجم پودر هسته خرما در نمونه شاهد، تفکیک‌پذیری افزایش می‌یابد.



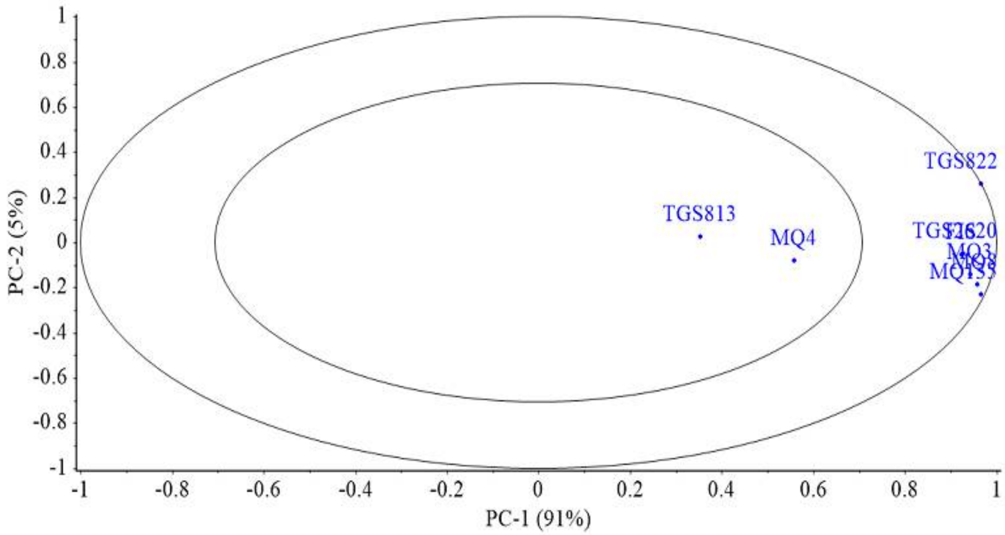
الف



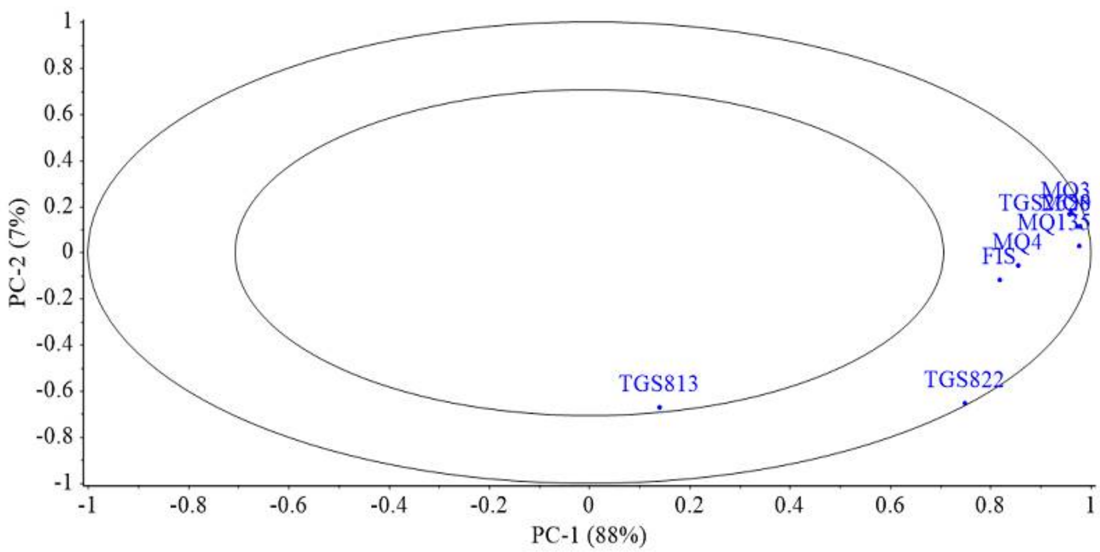
ب

شکل 2- نمودار اسکور PCA برای نمونه‌های مختلف فلفل سیاه. الف) براساس منشاء جغرافیایی ب) تقلب کبابه چینی (a1,2,3) ۱۰، ۲۰ و 30%، تقلب هسته خرما (b1,2,3) 10، 20 و 30%.

شکل 3- الف نمودار لودینگ برای دو مولفه اصلی نمونه‌های مختلف فلفل سیاه را نشان می‌دهد. هرچه مقدار لودینگ حسگری روی یکی از مولفه‌های اصلی بیشتر باشد (نزدیک به دایره بیرونی) نشان‌دهنده آن است که نقش بیشتری در تشخیص و تمایز بین نمونه‌های مورد نظر دارد و می‌توان برای کاهش هزینه‌های ساخت و پیچیدگی فرایند تشخیص الگو، حسگر یا حسگرهایی که کمترین تاثیر را در فرایند تشخیص داده دارند از آرایه حسگری حذف کرد.



الف



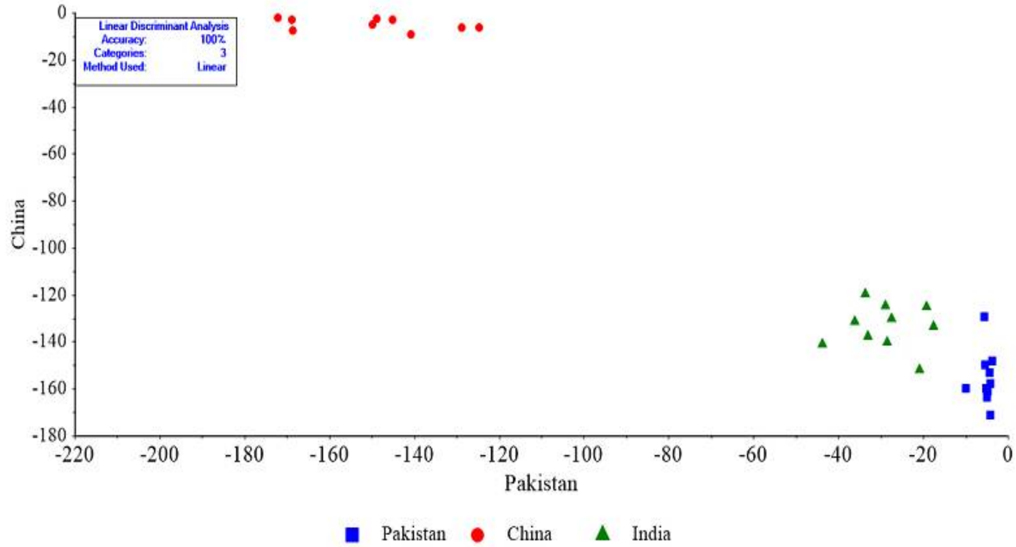
ب

شکل 3- نمودار لودینگ PCA برای نمونه‌های مختلف فلفل سیاه. الف) براساس منشاء جغرافیایی ب) تقلب کبابه چینی و پودر هسته خرما

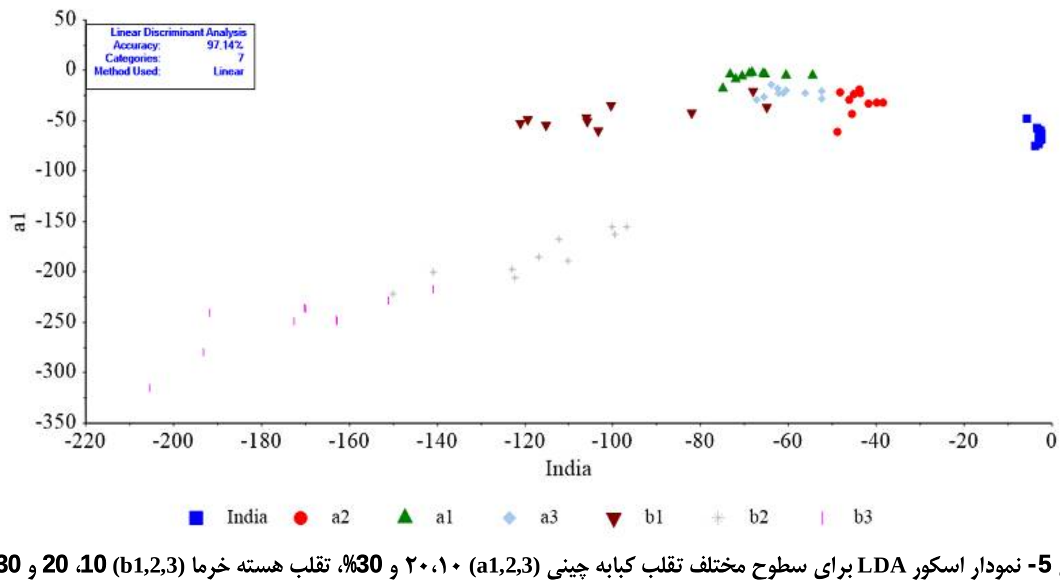
مطابق شکل 3- الف حسگرهای TGS822 و MQ135 و MQ8 بالاترین مقدار ضرایب لودینگ را دارند. MQ135 و MQ8 با وجود این که ضرایب لودینگ بالایی دارند اما با توجه به نزدیکی مقادیر آن‌ها به یکدیگر، می‌توان نتیجه گرفت که این دو حسگر تقریبا تاثیر یکسانی در فرایند تشخیص الگو دارند و می‌توان برای سهولت فرایند تحلیل، فقط از یکی از آن‌ها استفاده کرد. حسگرهای MQ4 و TGS813 با کمترین ضرایب لودینگ کمترین تاثیر را در طبقه‌بندی نمونه‌های فلفل سیاه داشته و می‌توان به‌منظور کاهش هزینه‌ها آن‌ها را از آرایه حسگری حذف نمود. در شکل 3- ب نمودار لودینگ مولفه‌های اصلی تقلب‌های کبابه چینی و هسته خرما نشان داده شده است. حسگرهای MQ8، TGS2620، MQ3 و TGS822 بالاترین

مقدار ضرایب لودینگ را به‌خود اختصاص دادند. حسگر TGS813 نیز کمترین ضریب لودینگ را به خود اختصاص داده است.

تحلیل تفکیک خطی LDA نیز به‌منظور بررسی کارایی سامانه در تشخیص و تفکیک بین بوهای فرار متصاعد شده از نمونه‌ها استفاده شد. براساس شکل 4 دقت این روش در طبقه‌بندی نمونه‌های شاهد فلفل سیاه 100% و برای سطوح تقلب با توجه به شکل 5، 97/14% به‌دست آمد. در نمودار مربوط به نمونه‌های شاهد فلفل سیاه، نمونه‌های مختلف به‌خوبی از یکدیگر تفکیک شده‌اند. در تقلب کبابه چینی سطوح مختلف تقلب از یکدیگر و از نمونه شاهد تفکیک شده‌اند و در تقلب هسته خرما سطح تقلب 10% از سایر سطوح تفکیک شده است ولی سطوح تقلب 20 و 30% به میزان بسیار کمی همپوشانی دارند.



شکل 4- نمودار اسکور LDA برای نمونه‌های فلفل سیاه براساس منشاء جغرافیایی



شکل 5- نمودار اسکور LDA برای سطوح مختلف تقلب کبابه چینی (a1,2,3) ۱۰، ۲۰ و 30%، تقلب هسته خرما (b1,2,3) 10، 20 و 30%.

همچنین از ماشین‌بردار پشتیبان نوع Nu-SVM برای طبقه‌بندی داده‌ها استفاده شد، پارامترهای Nu و γ با سعی و خطا و با به حداقل رساندن خطای اعتبارسنجی تعیین شدند. در این تحلیل از چهار نوع تابع کرنل خطی، چندجمله‌ای، گاوسی و سیگموئیدی استفاده شد. نتایج حاصل از این روش در جدول 3 آورده شده است. نتایج نشان داد که SVM با تابع گاوسی بالاترین دقت را در طبقه‌بندی نمونه‌های فلفل سیاه براساس منشاءجغرافیایی دارد. برای این تابع پارامترهای

بهینه SVM، Nu=0/66 و γ=0/5 و همچنین دقت آموزش و دقت اعتبارسنجی در طبقه بندی برابر با 96/66% به‌دست آمد.

مطابق جدول 3 در طبقه‌بندی این دو تقلب تابع گاوسی بیشترین دقت را در طبقه‌بندی از خود نشان داد و برای این تابع Nu=0/5 و γ=0/5 و دقت آموزش و دقت اعتبارسنجی به‌ترتیب برابر 92/85 و 90% به‌دست آمد.

در این پژوهش از الگوریتم J48 استفاده شد. با توجه به ماهیت این الگوریتم و نیاز به داده‌های آموزشی و آزمایشی برای ارزیابی عملکرد درخت تصمیم از روش K-fold Cross validation استفاده شده است.

جدول 3- عملکرد و پارامترهای ماشین‌بردار پشتیبان

ادویه	تابع کرنل	Nu	γ	دقت طبقه‌بندی در آموزش (%)	دقت طبقه‌بندی در اعتبار سنجی (%)
فلفل سیاه	خطی	0/663	1	90	90
	چندجمله‌ای	0/663	0/599	90	90
	گاوسی	0/663	0/599	96/66	96/66
	سیگموئیدی	0/663	0/027	83/33	86/66
تقلب	خطی	0/554	1	88/57	85/71
	چندجمله‌ای	0/554	0/599	87/14	84/28
	گاوسی	0/554	0/599	92/85	90
	سیگموئیدی	0/554	0/599	14/28	25

در این روش به پارامتر K نیاز است و با توجه به مقدار این پارامتر، K مرتبه و در هر مرتبه $\frac{1}{K}$ داده‌ها، به‌عنوان مجموعه داده‌های آزمایشی برای ارزیابی مدلی که با بقیه داده‌ها ایجاد شده است استفاده می‌شود. در انتها نیز به‌عنوان خروجی نهایی، میانگین این K مرتبه محاسبه می‌شود. K= 10 در نظر گرفته شد.

جدول 4- ماتریس اغتشاش حاصل از درخت تصمیم‌گیری برای نمونه‌های فلفل سیاه براساس منشاء جغرافیایی

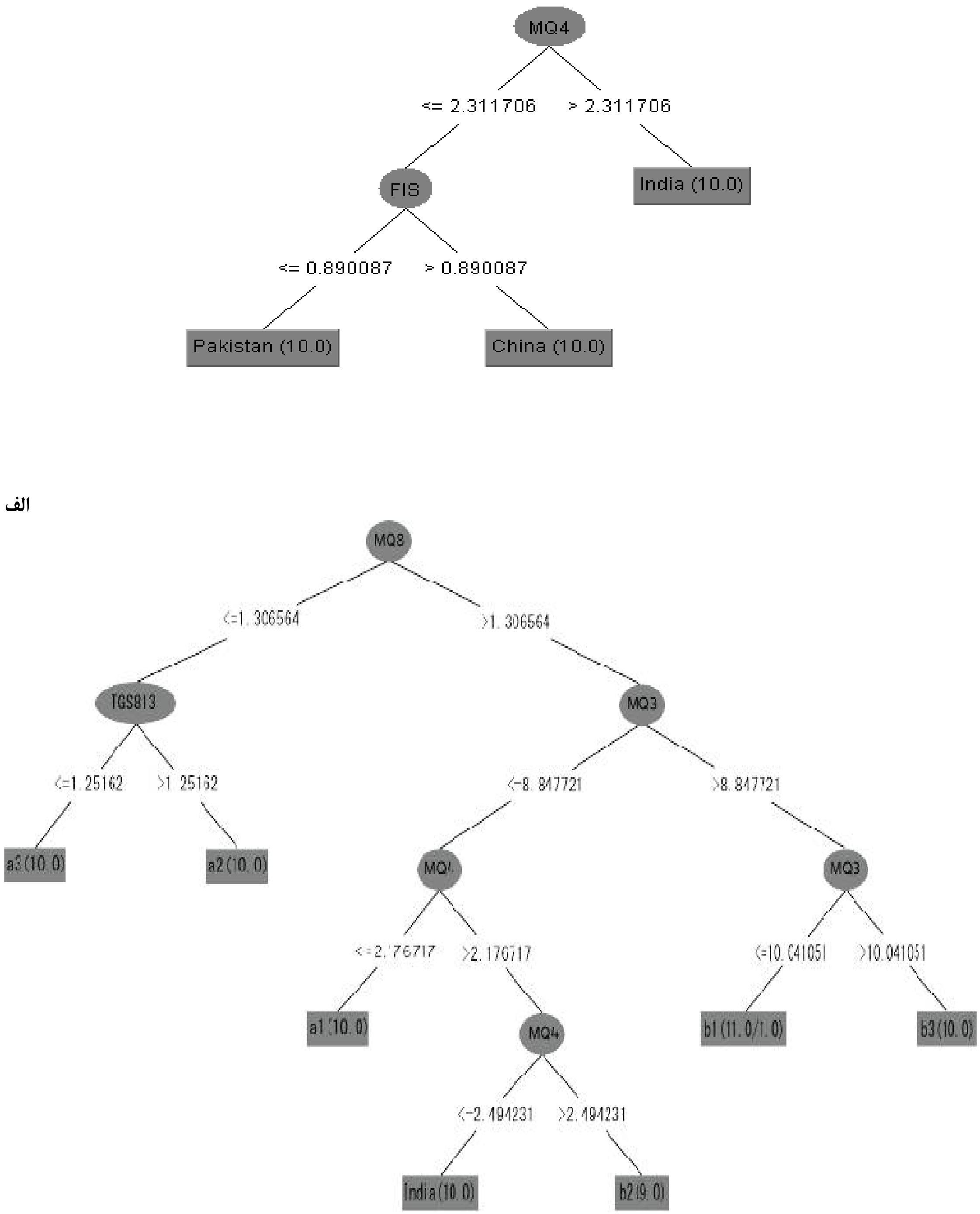
مشاهده/ پیش‌بینی	c	b	a
هندی	10	0	0
پاکستانی	1	0	9
چینی	0	10	0

جدول 5- ماتریس اغتشاش حاصل از درخت تصمیم‌گیری برای تقلب‌های کبابه چینی و هسته خرما

نمونه تقلب	مشاهده/ پیش‌بینی	g	f	e	d	c	b	a
	هندی	0	0	0	0	0	0	10
کبابه چینی	%10	0	0	1	1	0	8	0
	%20	0	0	0	1	9	0	0
	%30	0	0	0	9	1	0	0
هسته خرما	%10	1	0	9	0	0	0	0
	%20	0	7	2	0	0	0	1
	%30	10	0	0	0	0	0	0

نمودار گرافیکی نتایج درخت تصمیم در طبقه‌بندی داده‌ها در شکل 6 آورده شده است. یکی از موارد مهمی که در طبقه‌بندی داده‌های تاثیر مستقیم دارد، اندازه درخت‌های طبقه‌بندی می‌باشد. هرچه تعدا برگ‌ها و گره‌ها بیشتر باشد یا به عبارتی درخت بزرگتر شود، دقت آن در طبقه‌بندی کاهش می‌یابد. درخت تصمیم گیری به‌دست آمده برای نمونه فلفل سیاه از 3 برگ و 5 گره و برای تقلب‌های کبابه

چینی و هسته خرما از 7 برگ و 13 گره تشکیل شده است. این درخت‌ها به علت کم حجم بودن دقت مناسبی را ارائه داده‌اند.



شکل 6- نمودار درخت تصمیم در طبقه‌بندی نمونه‌های مختلف فلفل سیاه. الف) براساس منشاء جغرافیایی ب) تقلب کبابه چینی و تقلب هسته خرما

جدول 6- نتایج به‌دست آمده از درخت تصمیم‌گیری‌های ایجاد شده

ادویه	دسته	TP Rate	FP Rate	Precision	Recall	F-Measure
فلفل سیاه	هندی	1	0/05	0/909	1	0/902
	پاکستانی	0/9	0	1	0/9	0/947
	چینی	1	0	1	1	1
	میانگین وزنی	0/96	0/017	0/907	0/967	0/967
	هندی	1	0/017	0/909	1	0/952
تقلب کبابه چینی	%10	0/8	0	1	0/8	0/889
	%20	0/9	0/017	0/9	0/9	0/9
	%30	0/9	0/033	0/818	0/9	0/857
تقلب هسته خرما	%10	0/9	0/05	0/75	0/9	0/818
	%20	0/7	0	1	0/7	0/824
	%30	1	0/017	0/909	1	0/952
	میانگین وزنی	0/886	0/019	0/898	0/886	0/885

ماتریس اغتشاش و نتایج محاسبه پارامترها برای نمونه‌های فلفل سیاه و تقلب‌های کبابه چینی و هسته خرما در جدول‌های 4، 5 و 6 آورده شده است. ماتریس اغتشاش حاصل از درخت تصمیم ارتباط بین کلاس‌های واقعی و پیش‌بینی شده را بیان می‌کند. پارامترهای ذکر شده با استفاده از این ماتریس به‌دست می‌آیند. نسبت داده‌های هر کلاس به کل جمعیت داده‌های موجود را وزن هر کلاس گویند و با محاسبه میانگین وزنی درخت می‌توان پارامترهای ذکر شده را برای کل درخت به‌دست آورد. با توجه به نتایج کلی درخت تصمیم‌گیری قادر به طبقه‌بندی نمونه‌های فلفل سیاه با دقت 96/66% و سطوح مختلف تقلب با دقت 88/5% می‌باشد.

با توجه به جدول 6 مقدار پارامترهای F-Measure و TP Rate برای نمونه نمونه پاکستانی در ادویه فلفل سیاه و سطوح 10% در تقلب کبابه چینی و 20% درتقلب هسته خرما کمتر از سایر نمونه‌ها و سطوح دیگر بوده است. همچنین مقدار پارامتر F-Measure در سطح 30% در تقلب کبابه چینی و سطح 10% در تقلب هسته خرما کمتر از سایر سطوح می‌باشد. این امر نشان دهنده خطای بیشتر در دسته‌بندی داده‌های مربوط به این کلاس‌ها می‌باشد.

از فناوری ماشین بویایی به‌طور موفقیت‌آمیزی در طبقه‌بندی و تشخیص انواع مواد غذایی و کشاورزی استفاده شده است. اما در زمینه تشخیص و طبقه‌بندی انواع مختلف گیاهان دارویی و ادویه‌جات تحقیقات زیادی صورت نگرفته است. در این تحقیقات اندک صورت گرفته از سامانه ماشین بویایی براساس حسگرهای MOS استفاده شده است. طهری و همکاران (2016) از یک سامانه ماشین بویایی به‌منظور طبقه‌بندی زیره سبز براساس پنج منشاء جغرافیایی و تشخیص تقلب پودر گشنیز در پودر زیره سبز استفاده کردند. عملکرد این سامانه در طبقه‌بندی و تفکیک نمونه‌های بدون تقلب از سطوح تقلب با توجه به نتایج حاصل از روش‌های تفکیک خطی، خوشه‌بندی سلسله مراتبی و ماشین‌بردار پشتیبان، غیررضایت بخش گزارش شد. کارمونا و همکاران (2006) به‌منظور بررسی منشاء جغرافیایی زعفران (ایران، مراکش، یونان و اسپانیا) از یک سامانه ماشین بویایی در ترکیب با روش TD-GS-MS استفاده کردند. این سامانه با استفاده از تحلیل مولفه‌های اصلی قادر به تفکیک نمونه‌ها از منشاء جغرافیایی‌های متفاوت با دقت 90% بود. مقایسه نتایج به‌دست آمده از این پژوهش با استفاده از روش‌های PCA، LDA، SVM و DT با سایر پژهش‌های پیشین، نتایج کار را کاملا رضایت‌بخش نشان می‌دهد.

نتیجه گیری

در این پژوهش از یک سامانه ماشین بویایی بر پایه هشت حسگر نیمه‌هادی اکسید فلزی در ترکیب با روش‌های تشخیص الگو به‌منظور تشخیص نمونه‌های مختلف فلفل سیاه براساس منشاء جغرافیایی و تشخیص تقلب‌های کبابه چینی و پودر هسته خرما استفاده شد. در مقایسه تحلیل مولفه‌های اصلی (PCA)، تحلیل تفکیک خطی (LDA)، ماشین بردار پشتیبان (SVM) و درخت تصمیم‌گیری (DT) برای طبقه‌بندی نمونه‌های فلفل سیاه، روش LDA با دقت 100% برای نمونه‌های فلفل سیاه براساس منشاء جغرافیایی و دقت 97/14% برای تقلب‌ها بالاترین دقت طبقه‌بندی را نسبت به سایر روش‌های مورداستفاده را دارا می‌باشند. همچنین با انجام دادن آنالیز لودینگ، قابلیت حسگرها محاسبه شده و مشخص شد که آرایه حسگری مورد نظر برای تشخیص اصالت و تقلب در فلفل سیاه قابل قبول است. آگاهی داشتن از نوع حسگرها و توانایی آرایه حسگری ما

را در تصمیم‌گیری مناسب برای انتخاب، تغییر و ساخت حسگرهای مربوط به اهداف تعیین شده و تعیین زمان‌بندی‌های مختلف کمک می‌کند. نتایج به‌دست آمده بیانگر این موضوع است که سامانه ماشین بویایی در ترکیب باروش‌های تشخیص الگو توانایی تشخیص و طبقه‌بندی نمونه‌های مختلف فلفل سیاه و تقلبات کبابه چینی و هسته خرما را دارا می‌باشد. با توجه به توانایی سامانه می‌توان پیشنهاد کرد که تحقیقات بیشتری در طبقه‌بندی و تشخیص تقلب دیگر مواد غذایی و کشاورزی انجام شود.

منابع

امیدبیگی، ر.، 1384، تولید و فرآوری گیاهان دارویی، انتشارات آستان قدس رضوی.

توحیدی، م.؛ قاسمی ورنامخواستی، م.؛ غفاری نیا، و.؛ محتسبی، س.س. و بنیادیان، م.، 1395، ساخت و توسعه یک سامانه ماشین بویایی در ترکیب با روش‌های شناسایی الگو برای تشخیص تقلب فرمالین در شیر خام، مهندسی بیوسیستم ایران، دوره 47 ، شماره 4، ص 1-10.

حاجی نژاد، م.، محتسبی، س. س.، قاسمی ورنامخواستی، م. و آغباشلو، م.، 1396، تشخیص تقلب در نمونه عسل کنار با استفاده از یک سامانه ماشین بویایی، ماشین‌های کشاورزی، جلد 7، شماره 2، ص 450-439.

سیدی، ع. و عابدینی، م.، 1390، شیمی عمومی با نگرش کاربردی جلد 2، ویرایش اول، چاپ هفتم.

طلوعی اشلقی، ع.؛ پورابراهیمی، ع.؛ ابراهیمی، م. و قاسم احمد، ل.، 1391، پیش‌بینی عود مجدد سرطان پستان به کمک سه تکنیک داده کاوی، بیماری‌های پستان ایران، سال 5، شمارع 4، ص 23-34.

قاسمی ورنامخواستی، م.، 2011، طراحی و توسعه و پیاده‌سازی یک سیستم زبان الکترونیک بر پایه حسگرهای نیمه‌هادی اکسید فلزی (MOS) در ترکیب با روش‌های تجزیه و تحلیل تشخیص الگو برای بررسی تشخیص کیفیت و تغییرات آبجو، پایان‌نامه دکتری، دانشگاه تهران.

کیانی، س. و مینایی، س.، 1394، امکان سنجی استفاده از سامانه‌های هوشمند بر پایه ماشین بویایی و ماشین بینایی به منظور بررسی کیفیت و تعیین مواد موثر فرآورده‌های گیاهان داروئی (مطالعه موردی زعفران)، همایش ملی گیاهان دارویی و داروهای گیاهی.

Arshak, K.E., Moore, G. M., Lyons, J., Harris, S. & Clifford., 2004, A review of gas sensors employed in electronic nose applications, *Sensor review*, 24, 181-198.

Baldwin, E.A., Bai, J., Plotto, A. & Dea, S., 2011, Electronic Noses and Tongues: Applications for the Food and Pharmaceutical Industries, *Sensors*, 11(5), 4744-4766.

Banach, U., Tiebe, C. & Hubert, T, 2012, Multigas Sensors for the Quality Control of Spice Mixtures, *Food Control*, 26, 23-27.

Banerjee, D., Chowdhary, S., Chakraborty, S. & Bhattacharyya, R., 2017, Recent Advances in Detection of Food Adulteration. *Academic Press*, 145-160.

Bhattacharyya, N., Bandyopadhyay, R., Bhuyan, M., Tudu, B., Ghosh, D. & Jana, A., 2008, Electronic nose for black tea classification and correlation of measurements with “Tea Taster” marks, IEEE transactions on instrumentation and measurement, 57(7), 1313-1321.

Carmona, M., Martínez, J., Zalacain, A., Rodríguez-Méndez, M. L., de Saja, J. A. & Alonso, G. L., 2006, Analysis of saffron volatile fraction by TD-GC–MS and e-nose, *European Food Research and Technology*, 223, 96

Gliszczyńska-Świgło, A. & Chmielewski, J., 2017, Electronic nose as a tool for monitoring the authenticity of food, *Food Analytical Methods*, 10(6), 1800-1816.

Heidarbeigi, K., Mohtasebi, S.S., Foroughirad, A., Ghasemi-Varnamkhasti, M., Rafiee, S. & Rezaei, K, 2015, Detection of adulteration in saffron samples using electronic nose, *International Journal of Food Properties*, 18(7), 1391-1401.

Kiani, S., Minaei, S. & Ghasemi-Varnamkhasti., M., 2016a, Application of Electronic Nose Systems for Assessing Quality of Medicinal and Aromatic Plant Products: A review, *Journal of Applied Research on Medicinal and Aromatic Plants*, 3(1), 1-9.

Kiani, S., Minaei, S. & Ghasemi-Varnamkhasti, M., 2016b, Portable Electronic Nose as an Expert System for Aroma-based Classification of Saffron, *Chemometrics and Intelligent Laboratory Systems*, 156, 148-156.

Li, C.P., Heinemann, R. & Sherry., 2007, Neural network and Bayesian network fusion models to fuse electronic nose and surface acoustic wave sensor data for apple defect detection, Sensors and Actuators B: *Chemica*l, 125, 301-310.

Li, Ch., Xu, F., Cao, Ch., Shang, M.Y., Zhang, C.Y., Yu, J., Liu, G.X., Wang, X. & Cai, SH.C., 2013, Comparative analysis of two species of Asari Radix et Rhizoma by electronic nose, headspace GC–MS and chemometrics, *Journal of Pharmaceutical and Biomedical Analysis*, 85, 231-238.

Liu, H., Zeng, F.K., Wang, Q.H. & Wu, H. S., 2013, Studies on the chemical and flavor qualities of white pepper (*Piper nigrum* L.) derived from five new genotypes, *European Food Research and Technology*, 237(2), 245-251.

Lozano, J., Santos, J.P. & Horrillo, M.C., 2005, Classification of white wine aromas with an electronic nose, *Talanta*., 67, 610–616.

Mahmoudi, E, 2009, Electronic nose technology and its applications, *Sensors & Transducers*, 107, 17.
Patil, T.R. & Sherekar, S.S., 2013, Performance analysis of Naive Bayes and J48 classification algorithm for data classification, International Journal of Computer Science and Applications, 6(2), 256-261.
Pearce T.C., Gardner J.W., Friel S., Barlett P.N. & Blair., N., 2003, Electronic nose for monitoring the flavor of beers, Analyst, 118, 371–377.
Peter, K.V., (Ed.), 2012, Handbook of herbs and spices, Elsevier.
Sanaeifar, A., Mohtasebi, S.S., Ghasemi-Varnamkhasti, M., Ahmadi, H. & Lozano, J., 2014, Development and application of a new low cost electronic nose for the ripeness monitoring of banana using computational techniques (PCA, LDA, SIMCA, and SVM), Czech Journal of Food Sciences, 32(6), 538– 548.
Sanaeifar, A., Mohtasebi, S.S., Ghasemi-Varnamkhasti, M.. & Ahmadi, H., 2015b, Design, development and implementation of a metal oxide semiconductor (MOS) based machine olfaction system for monitoring of banana ripeness, .Journal of Agricultural Machinery, 5(1), 111-121.
Shafiqul Islam, A.K.M., Ismail, Z., Saad, B., Othman, A.R., Ahmad, M.N. & Shakaff, A.Y.Md., 2006, Correlation studies between electronic nose response and headspace volatiles of *Eurycoma longifolia* extracts, Sensors and Actuators B, 120, 245-251.
Tahri, K., Tiebe, C., El Bari, N., Hübert, T. & Bouchikhi, B., 2016, Geographical provenience differentiation and adulteration detection of cumin by means of electronic sensing systems and SPME-GC-MS in combination with different chemometric approaches, Analytical Methods, 8(42), 7638-7649.
Tudu, B., Kow, B., Bhattacharyya, N. & Bandyopadhyay, R., 2008, November, Comparison of multivariate normalization techniques as applied to electronic nose based pattern classification for black tea, In Sensing Technology, 2008. ICST 2008. 3rd International Conference on (pp. 254-258). IEEE.

Iranian Food Science and Technology Research Journal
Vol. 16, No. 4, Oct- Nov 2020, p. 479-491

نشریه پژوهشهای علوم و صنایع غذایی ایران
جلد 16، شماره 4، مهر - آبان 1399، ص. 491-479

Implementation of an olfactory machine system for the classification of different types of black pepper based on geographical origin and detection of cheating in Indian black pepper

F. Jamalizadeh[1], M Ghasemi-Varnamkhasti[2*], M. Ghasemi Nafchi$_3$, M. Tohidi[4], M. Dowlati[5]

Received: 2018.11.06
Accepted: 2019.09.08

Introduction: Spices are the most valuable medicinal plants used in food and medical science industries and due to quality and price diversity between various species, distinction, classification and separation of them based on purity and quality degree have great importance. Spices are produced in different countries, including India, Pakistan, China, and East and South Asian countries. The difference in the percentage of aromatic compounds in various types of spices from different regions has led to a distinction between spices. Also, profitable individuals for economic purposes and more profit without regard to the general health of the community will lead to the creation of adulteration in different types of spices. The most important of these adulterations is the addition of volatile ingredients such as cubeb pepper and palm kernel powder in black pepper.

Materials and Methods: In this study, an olfactory machine system based on eight metal oxide semiconductor sensors in combination with pattern recognition methods were used to classify and separate of black pepper samples based on geographic origin and also to detect cubeb pepper adulteration and palm kernel powder. The adulterated black pepper samples were tested with different adulteration levels (10, 20 and 30%).The fractional method was used to improve and optimize the electronic nose output signals before entering diagnostic methods. In order to analyze the extracted data from the sensor response signal, the principal component analysis method (PCA) was used. Based on the results, PCA with two main components of 96% for black pepper and 95% of cubeb pepper and palm kernel adulteration can be described from the variance of data. Also, three methods of linear separation analysis (LDA), Support vector machine (SVM) and decision tree (DT) were used to classify the samples. The use of the LDA method for black pepper showed a classification precision of 100%, and for adulterations, accuracy was 97.14%.

Results and Discussion: The results showed that SVM with Gaussian function has the highest accuracy in classifying black pepper samples, cubeb pepper, and palm kernel adulteration Also, the success rate of the DT method in separating and categorizing black pepper, cubeb pepper, and palm kernel was 96.66% and 88.5% respectively.

According to the results obtained, the machine olfaction system in combination with pattern recognition methods has the ability to detect and classify different black pepper samples from different geographical origin and the lowest level of adulteration.

Keywords: Nose electronics, Black pepper, Cubeb pepper, Palm kernel.

1. MSc student, Department of Mechanical Engineering of Biosystems, Faculty of Agriculture, Shahrekord University.
2. Associate Professor, Department of Mechanical Engineering of Biosystems, Faculty of Agriculture, Shahrekord University.
3. Assistant Professor, Department of Horticultural Science Engineering, Faculty of Agriculture, Shahrekord University.
4. Ph.D., Department of Mechanical Engineering of Biosystems, Faculty of Agriculture, Shahrekord University.
5. Assistant Professor, Tuyserkan Faculty of Engineering and Natural Resources, Bu- Ali Sina University.
(*- Corresponding Author Email: ghasemymahdi@ut.ac.ir)